ON SOME MATRIX TRACE INEQUALITIES

Peter Renaud

Department of Mathematics & Statistics
University of Canterbury, Private Bag 4800
Christchurch, New Zealand

No. 165 July, 1998

Abstract: A number of authors have recently considered matrix trace inequalities of Cauchy-Schwartz type. We extend some of these results and look at some other classical inequalities which can be reformulated as more general trace inequalities. The main step in generalizing these, is to overcome non commutativity of matrix multiplication. In many cases this is not too difficult. The result is a more general class of inequalities, often easier to fomulate and no more difficult to derive.

ON SOME MATRIX TRACE INEQUALITIES

Peter Renaud

Department of Mathematics and Statistics
University of Canterbury
Private Bag 4800
Christchurch, New Zealand

ABSTRACT: A number of authors have recently considered matrix trace inequalities of Cauchy-Schwartz type. We extend some of these results and look at some other classical inequalities which can be reformulated as more general trace inequalities. The main step in generalizing these, is to overcome non commutativity of matrix multiplication. In many cases this is not too difficult. The result is a more general class of inequalities, often easier to formulate and no more difficult to derive.

1. CAUCHY-SCHWARTZ TYPE INEQUALITIES.

The classical Cauchy-Schwartz inequality can be stated quite generally for sequences as

$$\left|\sum w_n a_n b_n\right|^2 \leq \sum w_n |a_n|^2 \sum w_n |b_n|^2$$

where (w_n) is a weight sequence satisfying $w_n \geq 0, \sum w_n = 1$.

Recently a matrix version, which can be regarded as a non-commutative generalization, has been studied by a number of authors. (See e.g. Bellman [2], Yang [5].) While they only consider the unweighted case, their results can be summarized as

THEOREM 1. *Let W be a trace matrix (i.e. W is positive semi-definite with $tr(W) = 1$). Then for matrices A, B of suitable dimension,*

$$|tr(WA^*B)|^2 \leq tr(WA^*A)tr(WB^*B).$$

This result is well known and depends only on the observation that

$$\langle A, B\rangle = tr(WA^*B)$$

defines an inner product.

Note that if all matrices commute then, at least in the Hermitian case, they can be simultaneously diagonalized and we are reduced to the classical case. For this reason, Theorem 1 can be regarded as a non-commutative version.

(The fact that the product of two positive definite matrices has positive trace is of course well known. What is more interesting and less well known is that such a product, though usually not Hermitian, must still have positive eigenvalues. See Coope [4].)

There is an allied inequality due to Bohr, which many authors treat as different to that of Cauchy-Schwartz, but which in fact is easily derivable from the latter.

BOHR'S INEQUALITY. If $z_1, z_2 \in \mathbf{C}$, $c > 0$ then

$$|z_1 + z_2|^2 \le (1+c)|z_1|^2 + \left(1 + \tfrac{1}{c}\right)|z_2|^2.$$

This has been generalized by Archbold [1] as follows:

GENERALIZED BOHR INEQUALITY. If $z_1, z_2, \ldots, z_n \in \mathbf{C}$ and $a_1, a_2, \ldots, a_n > 0$ with $\sum_{k=1}^{n} \frac{1}{a_k} = 1$ then

$$|z_1 + z_2 + \cdots + z_n|^2 \le a_1|z_1|^2 + a_2|z_2|^2 + \cdots + a_n|z_n|^2.$$

For this inequality too, there is a matrix version.

THEOREM 2. *Let W be a positive definite trace matrix. Then for any matrix A,*

$$|\mathrm{Tr}(A)|^2 \le Tr\left(W^{-1}A^*A\right).$$

Proof. Put $W = I$ in Theorem 1 and write as

$$|Tr(B^*C)|^2 \le Tr(B^*B)Tr(C^*C),$$

Then given W, A, let $B = W^{\frac{1}{2}}, C = W^{-\frac{1}{2}}A^*$ and the result follows from simple properties of the trace.

There are a number of other classical inequalities which are easy consequences of the Cauchy-Schwartz inequality although again they are sometimes represented as independent. Typical is Walsh's inequality ([3] p. 66).

If $w_n \geq 0, a_n > 0$, then $\sum (w_n a_n) \sum \left(\frac{w_n}{a_n}\right) \geq \left(\sum w_n\right)^2$.

This generalizes to

THEOREM 3. *If W is positive and A is positive definite, then*

$$\operatorname{Tr}(WA)\operatorname{Tr}(WA^{-1}) \geq (\operatorname{Tr}(W))^2.$$

Proof. Clearly we lose no generality in assuming that W is a trace matrix. Replace A by $A^{\frac{1}{2}}$ (the usual positive definite square root) and B by $A^{-\frac{1}{2}}$ in the generalized Cauchy-Schwartz inequality.

2. POLYA-SZEGO AND KANTOROVICH INEQUALITIES.

If one assumes certain mild boundedness conditions on sequences, then a partial reverse Cauchy-Schwarz inequality is possible. The clearest example is the Polya-Szego inequality ([3] p. 208).

Suppose $0 < m \leq b_n, c_n \leq M$. then

$$\sum b_n^2 \sum c_n^2 \leq \frac{1}{4}\left(\frac{M}{m} + \frac{m}{M}\right)^2 \left(\sum b_n c_n\right)^2.$$

This is easily deduced from the Kantorovich inequality ([3] p. 201).

If $0 < m \leq a_n \leq M$, $w_n \geq 0$ and $\sum w_n = 1$, then

$$\left(\sum w_n a_n\right)\left(\sum \frac{w_n}{a_n}\right) \leq \frac{(M+m)^2}{4Mm}.$$

(For if b_n, c_n are given as above, then let $a_n = \frac{b_n}{c_n}$ and $w_n = \frac{b_n c_n}{\sum b_n c_n}$. Then $\frac{m}{M} \leq a_n \leq \frac{M}{m}$ and by the Kantorovich inequality

$$\frac{\sum b_n^2}{\sum b_n c_n} \frac{\sum c_n^2}{\sum b_n c_n} \leq \frac{\left(\frac{M}{m} + \frac{m}{M}\right)^2}{4}$$

which is the required result.)

These two inequalities also have matrix versions.

THEOREM 4 (Generalized Kantorovich inequality)

Let A be positive definite with $0 < mI \leq A \leq MI$ and let W be a trace matrix. Then

$$\mathrm{Tr}(WA)\,\mathrm{Tr}\big(WA^{-1}\big) \leq \frac{(M+m)^2}{4Mm}.$$

(Here, as usual, $A \geq B$ means that $A - B$ is positive semi definite.)

Proof. We have $M^{-1}I \leq A^{-1} \leq m^{-1}I$ so that $(MI - A)\big(m^{-1}I - A^{-1}\big)$ is non negative. So

$$\begin{aligned} 0 &\leq \mathrm{Tr}\big(W(MI - A)\big(m^{-1}I - A^{-1}\big)\big) \\ &= \frac{M}{m} - \frac{1}{m}\mathrm{Tr}(WA) - M\mathrm{Tr}\big(WA^{-1}\big) \end{aligned}$$

i.e. $\mathrm{Tr}(WA) + Mm\mathrm{Tr}\big(WA^{-1}\big) \leq M + m.$

But for positive numbers a, b we have $2\sqrt{Mmab} \leq a + Mmb$ so that $2\sqrt{Mm\mathrm{Tr}(WA)\mathrm{Tr}WA^{-1}} \leq M + m$ and the result follows.

The special case where W is the projection onto a vector $\mathbf{x}$ (which can always be assumed normalized) gives $\mathrm{Tr}(WA) = \langle A\mathbf{x}, \mathbf{x}\rangle$ and we recover the result

$$\langle A\mathbf{x}, \mathbf{x}\rangle\langle A^{-1}\mathbf{x}, \mathbf{x}\rangle \leq \frac{(\lambda + \mu)^2}{4\lambda\mu}$$

where λ and μ are respectively the smallest and largest eigenvalues of A.

In the same spirit, there is a generalized Polya-Szego inequality. Perhaps surprisingly it does not follow from the Kantorovich inequality as it does in the classical case, due to lack of commutativity. But this is easily overcome.

THEOREM 5. (Generalized Polya-Szego inequality)

Suppose that $0 < mI \leq A, B \leq MI$. *Then*

$$\operatorname{Tr}(A^2)\operatorname{Tr}(B^2) \leq \frac{1}{4}\left(\frac{M}{m} + \frac{m}{M}\right)^2 (\operatorname{Tr}(AB))^2.$$

Proof. $mA \leq mMI \leq MB$ and similarly $mB \leq MA$.

So

$$\begin{aligned} 0 &\leq \operatorname{Tr}((MB - mA)(MA - mB)) \\ &= (M^2 + m^2)\operatorname{Tr}(AB) - mM(\operatorname{Tr}(A^2) + \operatorname{Tr}(B^2).) \end{aligned}$$

$\therefore mM(\operatorname{Tr}(A^2) + \operatorname{Tr}(B^2)) \leq (M^2 + m^2)\operatorname{Tr}(AB)$.

But $\operatorname{Tr}(A^2) + \operatorname{Tr}(B^2) \leq \frac{1}{4}(\operatorname{Tr}(A^2) + \operatorname{Tr}(B^2))^2$ and the result follows.

Sometimes matrix inequalities, while apparently more general, are in fact just restatements. As an example, the classical Arithmetic-Geometric Mean inequality

$$\frac{1}{n}\sum \lambda_i \geq \left(\prod \lambda_i\right)^{\frac{1}{n}} \quad \text{if } \lambda_i \geq 0$$

might lead to a "generalized" version.

THEOREM 6. *(Generalized Arithmetic-Geometric Mean inequality) If* A, B *are positive* n *by* n *matrices, then*

$$\frac{1}{n}\operatorname{Tr}(AB) \geq (\det(AB))^{\frac{1}{n}}.$$

Proof. This is a direct consequence of the "special case" above. For if we write $X = A^{\frac{1}{2}}BA^{\frac{1}{2}}$ then $\operatorname{Tr}(AB) = Tr(X)$ and det $(AB) =$ det (X). So we need only prove that $\frac{1}{n}Tr(X) \geq (\det\,(X))^{\frac{1}{n}}$. But assuming, as we may, that X is diagonal, the result is clearly equivalent to the classical Arithmetic-Geometric Mean inequality.

3. REARRANGEMENTS.

Let (a_k), $k = 1, \ldots, n$ be a positive sequence, and (b_k) be a rearrangement. The following are well known. (See e.g. [3] p. 20.)

$$(1) \sum_{k=1}^{n} \left(\frac{a_k}{b_k}\right) \geq n$$

$$(2) \sum_{k=1}^{n} (a_k b_k) \leq \sum_{k=1}^{n} a_k^2$$

Since a unitary transformation performs much the same effect on a matrix, e.g. it may permute its eigenvalues, we are led to the following

THEOREM 7. *Let A be n by n and positive definite and let U be unitary. Let $B = U^{-1}AU$.Then*

$$(1) \; Tr\left(B^{-1}A\right) \geq n$$

$$(2) \; Tr(AB) \leq Tr\left(A^2\right).$$

Proof. (2) follows from expanding the inequality $0 \leq Tr\left((A-B)^2\right)$, while (1) follows from Theorem 6 which implies that

$$\frac{1}{n} Tr\left(B^{-1}A\right) \geq \left(\det\left(B^{-1}\right)\det(A)\right)^{\frac{1}{n}} = 1.$$

REFERENCES

1. J. W. Archbold, *Algebra*, London, 1958.

2. R. Bellman, Some inequalities for positive definite matrices, General Inequalities, 2, Proceedings, 2nd International Conference on General Inequalities, (E. F. Beckenbach, Ed.), pp. 89–90, Birkhauser Verlag, Basel, 1980.

3. P. S. Bullen, D. S. Mitrinovic and P. M. Vasic (Eds), *Means and their inequalities*, D. Reidel, Dordrecht, Holland, 1988.

4. I. D. Coope, Matrix trace inequalities, *J. Math. Anal. Appl.* **188** *(1994), 999–1001.*

5. Xinmin Yang, Some trace inequalities for operators, *Bull. Aust. Math. Soc.*(1995)